DOCUMENT RESUME

ED 171 796 TM 009 444

AUTHOR Smith, Laura Spooner
TITLE Measures of High School Students' Expository Writing: Direct and Indirect Strategies.
PUB DATE Apr 79
NOTE 33p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Content Analysis; Essays; *Essay Tests; Evaluation Criteria; *Expository Writing; *Informal Assessment; *Objective Tests; *Performance Tests; *Scoring; Senior High Schools; Test Items; Test Reliability; Test Validity; *Writing Skills
IDENTIFIERS *Holistic Evaluation

ABSTRACT

The relationship among writing assessment strategies which may be applicable to competency based testing was examined in a study involving 128 high school students in six English classes in grades 11 and 12. Each student wrote two essays of at least 200 words on topics designed to test expository, or explanatory writing, and completed an objective test of writing-related competencies. Two grading methods were used for the essay tests: holistic scoring; and analytic scoring based on essay focus, development, organization, supporting statements, paragraph construction, and mechanics. Results showed moderately positive relationships between scores on the objective test and scores derived from the two methods of rating the essays. There was a strong relationship across all analytic subscales used to rate essays. High correlations between analytic and holistic scores were found. Holistic rating scores may have been indirectly influenced by the uniformity of the structural characteristics of the essays, as indicated by the emergence of the mechanics and supporting statements subscales as predictors of holistic ratings. Sample essay topics, evaluation criteria, and instructions for raters are appended. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED171796

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Measures of High School Students' Expository Writing: Direct and Indirect Strategies

Laura Spooner Smith
Center for the Study of Evaluation
University of California, Los Angeles

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Laura Spooner Smith

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

Paper presented at the annual meeting of the
American Educational Research Association
San Francisco, California
April, 1979

IM009 414

As demand increases for competency based tests of students' basic academic skills, additional requirements for measures of writing proficiency also have surfaced. The need today is for measures of writing that not only are technically sound, but which also serve as meaningful, efficient indicators of clearly defined writing competencies. Additionally, the demand is for measures that carry clear implications for instructional planning.

The present study was undertaken in an effort to examine relationships among writing assessment strategies which are potentially responsive to requirements of competency based testing. Three alternate strategies to measure secondary students' expository writing were developed and administered. Two of the strategies, direct measures, involved collecting and rating students' writing samples. The distinction between the direct measure strategies lay in the response criteria by which the samples were judged. One form of criteria, an Analytic rating scale, required raters to assign scores to six different characteristics of the writing samples. The other form of criteria, an Impressionistic rating scale, yielded a single score on the quality of each essay as an example of exposition. The third stragegy, an indirect measure, was an objective test of writing-related competencies derived from the Analytic rating scale.

Background

The measurement of writing and writing-related competencies historically has presented unique technical and validation problems. From pioneer assessment efforts to current composition tests, measurement experts have contended with such recurring problems as fluctuating test reliability, implementation

of efficient scoring procedures, and examining the validity of indirect, i.e., objective, measures of writing.

In recent years, additional requirements for measures of writing have arisen along with increasing demand for competency based tests of basic skills. This situation has prompted renewed attention to a fundamental concern in writing assessment: How to identify and define measurement tasks that will serve as efficient and instructionally meaningfully indicators of writing competence. Identifying test tasks (or items) to measure a domain of learning is, of course, an intellectually difficult aspect of competency based assessment, especially for complex behaviors such as the production of written discourse. Millman's (1974) assertion that a performance domain should be defined"...by those facets and elements that make a difference in how the learner responds" leads naturally to the question, what facets make a difference?

The Response Criteria Issue

In the measurement of writing skills, one "facet" that clearly makes a difference is the criteria by which writing samples are judged. Although a variety of guided scoring procedures for sorting or ranking written pieces currently are in use, most can be classified as either "analytic" or "holistic."

Analytic scoring procedures presume that a piece of writing can be viewed as consisting of component, but not necessarily independent, parts which are worthy of individual scrutiny. The procedures require raters to assign points to each of several specified aspects of a composition and yield an estimate of overall quality of the writing product, as well as sub-

scores on separate elements of the writing sample. Analytic rating procedures aptly meet the requirements of competency based assessment programs which call for information on specific skill strengths and deficits.

In contrast, holistic rating procedures assume that "...each factor that makes up writing skill is related to all other factors and that one factor cannot be easily separated from others" (Office of the Los Angeles County Superintendent of Schools, 1977). Under holistic procedures, raters assign a single value to a piece of writing. A specific type of holistic procedure, impressionistic rating scales, also require assignment of a single value to a written product. Unlike most holistic approaches, impressionistic scales involve only a minimal rubric to guide the judgments of raters. The apparent advantage of holistic rating procedures in general, and impressionistic scales in particular, is that they tend to be an efficent and reliable direct measure of writing performance. The drawback is the lack of precise information on the attributes of writing to which raters attend. Whether analytic and holistic procedures provide substantially comparable estimates of writing competence is an important question, especially within the context of competency based assessment.

The Response Mode Issue

Another "facet" in the measurement of writing skills is the form of response elicited from examinees. Indeed, an important issue in writing assessment centers on the relative merits of two contending response modes: production of a piece of writing, i.e., "writing samples," and selecting a response from among given alternatives, i.e., "objective" tests of writing.

That there is any question regarding the most appropriate approach for assessing composition skills at first may appear puzzling. Clearly, writing samples represent a direct measure of writing performance and therefore possess prima facie validity. This, in fact, is the principal argument supplied by proponents of writing samples as a means of composition assessment. Those who favor more indirect methods, i.e., objective tests, however, have pointed to the usual unreliability of writing samples, notably the difficulty of ensuring reliable scoring procedures and the tendency of a writer's performance to fluctuate in quality over time and task. Problems of unreliability in writing samples have been the subject of recent research which has resulted in promising procedures aimed at enhancing measurement reliability. Nonetheless, a clear advantage of objective methods of assessment is the efficiency with which the measures can be administered and scored, an economy of especial importance when sizeable aggregates of students are to be tested.

Arguments of efficiency, though, have failed to impress many professionals in the discipline of English, who have voiced concern over the inherent lack of content validity of indirect approaches. In response to such objections, proponents of indirect measures have cited findings from studies (e.g., Godshalk, 1966) which reveal statistically significant correlations between performance on objective tests of writing and performance on actual writing samples.

The limitation of many correlational and predictive studies of objective tests of writing for competency based assessment, however, centers on the criterion against which the items are validated, notably the use of holistic scoring procedures to rate the criterion essays. Recall that hol-

istic scoring results in a single score on the overall quality of an essay; separate scores on different characteristics of an essay are not provided. Consequently, no inferences can be drawn about possible relationships between the classes of skill t pped by objective items and the classes of skill exhibited in the criterion essay. At best, one can conclude that various combinations of skill measured by objective items are related in some unspecified fashion to actual writing performance. What appears to be needed is information on the relationships between well-defined classes of objective items and equally well-defined writing production measures.

Method

Subjects, 128 eleventh and twelfth grade students in six English classes in the Los Angeles area, were randomly assigned within each class to treatment groups determined by the order in which the measures were administered. Attachment 1 depicts the three measurement strategies. Each subject wrote two essays of at least 200 words on topics designed to elicit expository writing and completed an objective test of writing-related competencies. Two raters were trained to employ an Analytic rating scale and two to use an Impressionistic rating scale. The writing samples were scored by both rater pairs, resulting in four total scores for each sample. Final study reliability of ratings on the Analytic total scale was .90, with rater reliabilities for the six Analytic subscales ranging from .84 to .95. Rater reliability for the Impressionistic scale was .87.

Measures

The direct measure writing task employed in the study consisted of two major components: the writing topics (and directions) and two forms of rating criteria. The topics, designed to elicit like-samples of student writing, were intended to promote writing within the discourse domain of exposition, that is writing "...that explains or clarifies a subject" (Brooks and Warren, 1961). Task attributes guiding development of the topics included discourse mode, rhetorical purpose, content limits, and intended audience. Copies of the topics are presented in Attachment 2.

One form of rating criteria, the Analytic scale, was designed to reflect state-of-the-art pedagogical precepts and practice in composition. Elements on the scale were derived from an analysis of conventionally recognized structural features of exposition as indicated in curriculum guides and textbooks at the secondary level. The final version of the Analytic scale yielded scores on six subscales corresponding to the following elements of writing: essay focus (main idea), organization, development, support, paragraphing, mechanics. The range of points for each of the Analytic subscales was four (high) to one (low). The rating rubric contained descriptions of essay characteristics for each of the four levels for all six subscales. Attachment 3 presents an abbreviated version of the Analytic scale and an example of one subscale rating rubric.

The second form of response criteria, the Impressionistic rating scale, required raters to make a judgment regarding the overall quality of each writing sample as an example of effective exposition. The rating rubric directed raters to assign each essay a single numerical score by employing

a six-point scale, with six (high) and one (low). In addition to the scale, the rubric contained several definitions outlining prominent conventionally recognized features of exposition. A copy of the Impressionistic rubric is contained in Attachment 4.

Additionally, an indirect measure, a 37-item multiple-choice test, was developed to measure skills presumably related to actual production of expository writing. The skills covered in the test were identified through two related analyses. The first analysis consisted of a review of expository writing skills frequently emphasized in secondary composition curricula and instructional materials for which selected-response type practice was provided. For example, a typical exercise required students to identify from a list those details which either do or do not support a given generalization. The skills identified through the review were then arrayed against elements listed in the Analytic rating scale in order to determine which skills were conceptually analogous to the rating scale elements.

The final version of the Objective measure contained a subtest for each of the following Analytic scale elements: focus (main idea), development, organization, support, paragraphing. Each of the five subtests contained five similar-format items which were generated according to a set of test item specifications. Objectives for the subtests are presented in Attachment 5. The method by which items for a sixth subtest were generated was different from that of the first five subtests. Unlike the stimulus passages of the first five subtests, which were generated specifically for the objective test, the passages in subtest six

were drawn from actual samples of students' writing. The passages selected exhibited one or more of several types of errors, e.g, failure to state or imply the main idea, lack of supporting statements. Each passage was followed by four items directing the student to identify the statement(s) which exhibited a specified category of error.

Summary of Findings

Relationships Among Analytic Scale And Impressionistic Scale Scores

Correlations among scores from the two rating scales are presented in Attachment 6. As shown, scores on the six subscales comprising the Analytic scale proved to be highly related, with correlations ranging from .69 to .90. Correlations between subscale scores and the Analytic scale total scores ranged from .82 to .96.

The correlation between the Impressionistic scale scores and Analytic total scale scores (.81) indicated a strong association between the two rating strategies. The association extended to relationships between the six Analytic subscales and Impressionistic ratings, with correlations ranging from .65 to .80.

To further examine the relationship between Impressionistic and Analytic scores, Impressionistic scores were regressed on the six Analytic subscales, which jointly accounted for approximately 75 percent of the variation ($F=53.736$, $df=6,105$, $p < .01$) in Impressionistic scores. Two Analytic subscales, Mechanics ($F=30.789$) and Support ($F=18.365$), proved to be significant predictors to Impressionistic scale scores in the model with all six subscales (see Attachment 7).

The relatively strong associations among the Analytic subscales suggested that the relative importance of the subscales as predictors may be masked in the regression analyses. To examine this, a new composite variable, Structure, which was comprised of the sum of scores on four Analytic subscales (Organization, Focus, Paragraphing, Development) was entered into the equation. The combination of the three subscales, Mechanics, Support and Structure, accounted for 74 percent of the variation in Impressionistic scale scores ($F=105.159$, $df=3,108$, $p < .01$). Again, Support ($F=28.502$) and Mechanics ($F=31.468$) emerged as significant predictors to Impressionistic scores (see Attachment 7).

Relationships Among Rating Scale Scores and the Objective Measure

The correlation between the Objective measure total scores and the Analytic scale total scores was .61, while the correlation with Impressionistic scores was .65 (see Attachment 8). Correlations between the six Objective subtest scores and total scores on the two rating scales ranged from .55 to .23.

To examine the predictive relationship between the Objective measure and the two rating scales, the Impressionistic and Analytic scale total scores were independently regressed on the six major Objective subtests (see Attachment 9). Results of the regression analyses for the Impressionistic scale ($F=12.853$, $df=6,92$, $p < .01$) and for the Analytic scale ($F=10.338$, $df=6,92$, $p < .01$) indicated that the six Objective subtests jointly accounted for approximately 46 percent of the variation in Impressionistic scale scores

and for approximately 40 percent of the variation in Analytic scale scores. Two Objective subtests, Paragraphing (Subtest 4) and Paragraph Analysis (Subtest 6), proved to be significant predictors to both Impressionistic and Analytic scale total scores.

In an additional series of analyses, Analytic subscales were regressed on Analogous groups of Objective items. These calculations (reported in Attachment 10) resulted in significant F ratios for eight of the nine groups of Objective items.

Discussion

An interesting finding to emerge from the study concerned the pattern of strong relationships across the Analytic subscales. On the basis of these data alone, it is tempting to infer that the Analytic scale actually tapped a single unitary dimension of writing. Such an inference, however, overlooks an important facet of the writing task which may have affected the results -- the writing topic and directions.

The topics and directions of the study were specifically designed to elicit writing developed through a logically arranged structure of generalizations supported by specifics. Directions accompanying the topics prompted students in the following way:

> Remember, the purpose of your essay is to give an informative explanation...Back up your ideas with specific support, such as examples, facts, and other details. Make sure your essay is well-organized.

As anticipated, the topics and directions promoted uniformity in the rhetorical structures of the majority of students' writing samples, of course, with varying degrees in quality of execution. Whether differential sub-

scale scores will emerge when writing samples display more varied structural patterns is an area worthy of additional inquiry.

It is also possible that Impressionistic rating scores were indirectly a result of the relative uniformity of the structural characteristics of the essays. Once Impressionistic raters became habituated to the structural patterns of the majority of writing samples (through practice and training), they may have attended to a few prominent features -- other than structure -- which most noticeably discriminated among the essays.

The preceding notion is supported by the emergence of the two Analytic subscales, Mechanics and Support, as predictors to Impressionistic ratings. Judgments regarding a writer's command over mechanical aspects of writing can be made independently of the overall structure of a written product. Similarly, judgments regarding adequacy of support for generalizations embedded within an essay can be made without reference to the overall structure of a written piece. Given the relative conformity of structural patterns of the students' essays, Analytic and Impressionistic raters may have inadvertently attended to two readily discernible, and thus discriminating, features of the writing samples: mechanics and support.

This interpretation of statistical relations between the two rating strategies is not meant to imply that the discourse domain of exposition consists exclusively of two components, mechanics and support. The high correlations between Analytic scores and Impressionistic scores suggests that the Analytic subscales did, in fact, represent recognizable, if not necessarily independent, features of exposition. A provocative issue presents itself: Was the relative lack of independence among the Analytic

subscales -- and the high correlation with Impressionistic scores -- a function of the homogeneity of student responses? Or, are features of writing such as development, organization and main idea actually inseparable for purposes of rating?

Not too surprising were findings which revealed moderately positive relationships between Objective measure scores and those yielded by the two essay rating strategies. The positive relationship was expected, as the Objective items were designed to assess, at the levels of recognition and discrimination, those categories of skill measured by the rating strategies at the level of production. Moreover, previous studies have demonstrated that reasonably well-designed objective tests of writing invariably correlate with scores on writing samples, a phenomena which, commonsensically, can be accounted for by the global constructs of language (reading) ability or verbal aptitude.

Of more compelling interest to competency based test developers were the patterns of student performance on the Objective measure. Given that reading ability was likely to affect test performance, the majority of items were developed with the aim of minimizing reading difficulty, e.g., avoiding complex constructions, abstract content, advanced vocabulary. In fact, most of the items were designed to require students to make discriminations among individual sentences, rather than sentences embedded within prose passages.

The relatively high mean performance (see Attachment 11) of students on items requiring discrimination among individual sentences suggested that many of the students possessed the writing-related competencies being measured, such as selecting details to support generalizations, arranging given

ideas in a logical order, choosing statements to develop a given main idea. For many lower-ability students (as indicated by teacher ratings), though, these competencies were not expressed when stimulus competition within the task (e.g., number of words and sentences, sentence structure) was increased.

Especially worthy of consideration are properties of items within the two subtests (Subtests 4 and 6) which proved to be significant predictors to the essay-rating total scores. Both of these subtests required students to make discriminations among statements embedded within prose passages. Here again, the construct of verbal (or reading) ability is a convenient, but hardly satisfying, explanation for a statistical artifact.

The preceding discussion has highlighted some of the technical problems and issues associated with the measurement of writing and writing-related skills, but what about practical implications? Findings indicated that the two rating scales provided essentially comparable estimates of writing competence. By definition, analytic scales, such as the one employed, have a greater potential than impressionistic (or holistic) scales to paint a clear picture of students' writing strengths and deficits. As expected, the time required to train Analytic raters and to analytically rate the essays was slightly greater than the time required for Impressionistic procedures (approximately six hours). The expense, however, is likely to be outweighed by the usefulness of the information yielded by analytic procedures. The explicit nature of analytic scales provides instructional decision makers and students with clear information on the domain of learning being measured. Such information is likely to enhance dialogue among teachers, students and

administrators on the status of student writing and may serve as a basis for instructional planning, diagnosis and remediation.

The contribution of indirect measures of writing, however, raises a variety of issues requiring further study. One of the most basic issues centers on the nature of the relationship between direct and indirect measurement strategies. As demonstrated in the present study, as well as others, there exists an array of selected response tasks (beyond those measuring sentence level skills and writing mechanics) which are conceptually and statistically associated with writing production. If selected response tasks are to be useful within the context of competency based measurement, though, test developers must employ test items which are positively related to instructional efforts.

Sixteen years ago, Richard Braddock characterized research on composition as "...laced with dreams, prejudices, and makeshift operations" (Braddock, 1963). It probably is fair to say that the state of composition research has advanced, even accelerated, in recent years. This may be due in large part to growing acceptance of a research paradigm which views measurement and instruction as complementary pursuits. A continuing challenge to test developers, then, is to identify measurement tasks which, when practiced under appropriate instructional conditions, are likely to promote, not simply predict, acquisition of writing production skills.

REFERENCES

Braddock, R., Lloyd-Jones, R., & Schoer, L. Research in written composition. Champaign, Ill.: National Council of Teachers of English, 1963.

Brooks, C., & Warren, R. Modern rhetoric, shorter edition. New York: Harcourt, Brace & World, Inc., 1961.

Godshalk, F.I., Swineford, F., & Coffman, W. E. The measurement of writing ability. New York: College Entrance Examination Board, 1966.

Millman, J. Criterion-referenced measurement. In W. J. Popham (Ed.), Evaluation in education: Current applications. Berkeley, Ca.: McCurchan, 1974.

Office of the Los Angeles County Superintendent of Schools. A common ground for assessing competencies in written expression, review copy. Los Angeles: Division of Curriculum and Instructional Services, 1977.

ATTACHMENT 1

Measurement Strategies

Type of Strategy:

Direct Measure of Writing	Indirect Measure of Writing
1. Writing samples judged by Analytic scoring criteria.	3. Objective items requiring students to discriminate among given passages of written discourse.
2. Writing samples judged by Impressionistic scoring criteria.	

ATTACHMENT 2

Writing Topics

FINAL STUDY TOPICS

TOPIC I

Reading, writing, and arithmetic, the three R's, have become a serious problem for many young adults. In fact, every year thousands of American students graduate from high school unable to read simple words, write short essays, or do math. To understand this problem, experts have studied areas such as teaching methods, family life, and the effects of television. However, no one best explanation has been found.

Assume you have been asked by a group of parents and teachers to explain the problem. Write an essay in which you explain why many of today's students graduate from high school without basic skills. Your essay should focus on one (or more) of the "three R's", reading, writing or math.

Remember, the purpose of your essay is to give an informative explanation, not opinions. Back up your ideas with specific support, such as examples, facts, and other details. Make sure your essay is well-organized.

TOPIC 2

The use of alcohol, drugs, and cigarettes has become a serious problem for many young adults. In fact, every year the use of these substances hurts the lives of thousands of American teenagers by making them emotionally and physically ill. To better understand this problem, experts have studied areas such as family and school life, pressure from friends, and the effects of television and advertising. However, no one best explanation has been found.

Assume you have been asked by a group of parents and teachers to explain the problem. Write an essay to help explain why many of today's teenages abuse alcohol and drugs or smoke cigarettes. Your essay should focus on one (or more) of the substances, alcohol, drugs, or cigarettes.

Remember, the purpose of your essay is to give an informative explanation, not opinions. Back up your ideas with specific support, such as examples, facts, or other details. Make sure your essay is well-organized.

ATTACHMENT 3

Analytic Rating Scale (Abbreviated)

THE ANALYTIC RATING SCALE

Analytic Scale Elements

1. Essay Focus: The introduction or conclusion of the essay clearly indicates the subject and main idea of the whole essay.

2. Essay Development: All major subtopics ("main points") clearly relate to the main idea of the whole essay.

3. Essay Organization: The main idea is developed according to a clearly discernible method of organization.

4. Support: Generalizations and assertions are supported by specific clear supporting statements.

5. Paragraphing: The essay is composed of one or more clearly discernible units of thought, e.g., paragraphs.

6. Mechanics: The essay is free of intrusive mechanical errors.

Copyright: Laura Spooner Smith, 1979

ATTACHMENT 3 (con't.)

Analytic Rating Scale: Sample Rubric

ELEMENT 1

Essay Focus

The introduction (if deductively structured) or conclusion (if inductively structured) of the essay clearly indicates the subject and main idea of the whole essay.

4. The introduction (and/or conclusion) of this paper clearly conveys the main idea of the whole essay. It also limits the topic by alerting the reader to the key points covered in the body of the essay.

   Specifically, in the introduction (and/or conclusion):

   a. The subject of the essay is clearly identified.

   b. The main idea of the whole essay is clearly stated or implied.

   c. The topic is clearly limited. That is, key points (e.g., reasons, ideas) or major line(s) of reasoning treated in the essay are identified or summarized.

3. The introduction (and/or conclusion) of this paper conveys the main idea of the whole essay. It sets limits on the topic, but does not clearly suggest how the main idea is developed.

   Specifically, in the introduction (and/or conclusion):

   a. The subject of the essay is clearly identified.

   b. The main idea of the whole essay is clearly stated or implied.

   c. An attempt is made to limit the topic. That is, the number -- or type -- of key points is specified, but there is not clear reference to the substantive issues treated in the body of the essay.

2. The introduction (and/or conclusion) of this paper gives the reader a fairly clear sense of the main idea of the whole essay. However, neither the introduction nor the conclusion help focus -- or bring direction to -- the body of the paper.

   Specifically, in the introduction (and/or conclusion):

   a. The subject of the essay is identified.

   b. The main idea of the whole essay is stated or implied.

   c. No attempt is made to limit the topic.

ELEMENT 1 (continued)

Essay Focus

1. Neither the introduction nor the conclusion is helpful to the reader in obtaining any sense of the main idea of the essay.

   Specifically, in the introduction (and/or conclusion):

   a. The subject of the essay is not clearly identified or there is no reference to the subject

   AND/OR

   b. The main idea of the whole essay is not clearly stated or implied or no reference is made to the main idea, or the reference is confusing.

ATTACHMENT 4

Impressionistic Rating Scale

IMPRESSIONISTIC RATING SCALE RATER TRAINING MATERIALS

Impressionistic Rating Procedures

The purpose of Impressionistic Rating procedures is to form a single impression of a piece of writing as to how well it communicates a whole message to the reader. Impressionistic scoring assumes that each characteristic that makes up an essay -- organization of ideas, content, mechanics and so on -- is related to all other characteristics. Impressionistic scoring further assumes that the qualities of an essay cannot easily be separated from each other. In short, the procedure views a piece of writing as a total work, the whole of which is greater than the sum of its parts.

Discerning readers naturally will attend to, or be influenced by, some essay characteristics more than others. In Impressionistic scoring, however, readers are not expected to assign values to separate characteristics of an essay. Instead, they should arrive at a judgment regarding the overall quality of the essay.

For this rating session, you are being asked to form an overall impression concerning the effectiveness of the essays as examples of expository writing.

The Topics

You will be reading essays on two different topics. Both topics, though, were designed to elicit writing in the expository mode. Some views on exposition are given below.

.Exposition is the kind of discourse that explains or clarifies a subject.

.Exposition seeks to explain or inform through such methods as giving reasons or examples, comparing and contrasting, defining, ennumerating or through a combination of methods.

.Exposition explains why or how.

.Exposition promotes reader understanding of a subject,

Copyrignt: Laura Spooner Smith, 1979

Instructions to Readers

You are to read each essay quickly in order to form an overall impression of its quality. To assign the essay a score, consider the following question: To what extent is the essay an example of effective exposition?

Assign each paper a mark of 1 - 6 using the scale below.

High

6 = An excellent example of exposition

5 = A good example of exposition

Middle

4 = An adequate example of exposition

3 = A minimally adequate example of exposition

Low

2 = A poor example of exposition

1 = Barely readable and/or off the topic

PLEASE NOTE: The papers may not represent a "normal" distribution, that is, some categories may have very few papers while others may have many. In other words, do not intentionally sort an equal number of papers into all six categories.

ATTACHMENT 5

Objective Measure Subtest Objectives (Abbreviated)

I. Objective

(5 items) The student will be given a brief paragraph which is lacking either a topic or concluding sentence, and four alternate sentences. The student is to select the sentence which would serve as the most appropriate topic or concluding sentence, i.e., a statement of the paragraph's main idea.

II. Objective

(5 items) The student will be given a main idea statement for a multi-paragraph expository essay and alternate statements that might be included in the body of the essay. The student is to select the statement that most directly contributes to development of the given main idea.

III. Objective

(5 items) The student will be given a topic sentence for an expository paragraph and alternate statements that might be included in the paragraph as supporting detail. The student is to select the statement that does not provide specific support for the given topic sentence.

IV. Objective

(5 items) The student will be given a series of five to six lettered sentences which express two distinct thoughts (i.e., sub topics) which are related to the same overall main idea. The student is to indicate where one complete thought ends and another begins, i.e., where a new paragraph could logically begin.

V. Objective

(5 items) The student will be given five sentences which could be included in an expository essay: One statement of the essay's main idea; two "sub topic" sentences ("topic" or "concluding" sentences for individual paragraphs within the essay); two supporting details. The sentences will be given in scrambled order. The student is to indicate a logical order for the sentences.

ATTACHMENT 6

INTERCORRELATIONS AMONG IMPRESSIONISTIC AND ANALYTIC RATING SCALE SCORES

	Imp. Total	An. Total	I Focus	II Devel-opment	III Organ-ization	IV Sup-port	V Para-graph	VI Mech-anics
Impressionistic Total	1.00							
Analytic Total	.81	1.00						
Analytic Subscale:								
I. Focus	.65	.91	1.00					
II. Development	.69	.92	.90	1.00				
III. Organization	.76	.96	.89	.90	1.00			
IV. Support	.80	.86	.72	.71	.82	1.00		
V. Paragraphing	.74	.94	.80	.83	.91	.84	1.00	
VI. Mechanics	.79	.82	.69	.73	.72	.70	.70	1.00

ATTACHMENT 7

Regression of Impressionistic Scale on Analytic Scale

REGRESSION OF IMPRESSIONISTIC SCALE SCORES ON
SIX ANALYTIC SUBSCALES

Analytic Subscales	Unstandardized Coefficient	Standard Error of β	Standardized Coefficient	F
Mechanics	.585	.106	.424	30.789*
Focus	.233	.147	.196	2.537
Development	.046	.160	.040	.082
Organization	.293	.182	.265	2.492
Support	.514	.120	.424	18.365*
Paragraphing	.033	.148	.028	.048

df = 1,105

* = $p < .01$

$r^2 = .75$

R = .87

REGRESSION OF IMPRESSIONISTIC SCALE SCORES ON
THREE ANALYTIC SUBSCALES

Analytic Subscales	Unstandardized Coefficient	Standard Error of β	Standardized Coefficient	F
Mechanics	.582	.104	.422	31.468*
Support	.560	.105	.462	28.502*
Structure[1]	.017	.028	.056	.370

df = 1,108

* = $p < .01$

Structure[1] = Composite score of Organization, Focus, Development, Paragraphing

$r^2 = .74$

R = .86

ATTACHMENT 8

Correlations Among Rating Scales and Objective Measure

CORRELATIONS AMONG ANALYTIC AND IMPRESSIONISTIC TOTAL SCORES AND OBJECTIVE MEASURE SCORES

	Imp. Total	An. Total	I Main Idea	II Devel-opment	III Supt. Detail	IV Para-graph	V Organ-ization	VI Para. Anal-ysis	A Main Idea[1]	B Devel-opment[1]	C Organ-ization[1]	D Supt. Detail[1]	Total Test
Impressionistic Total	1.00												
Analytic Total	.81	1.00											
Objective Subtests													
I. Main Idea	.24	.23	1.00										
II. Development	.46	.37	.17	1.00									
III. Supt. Detail	.35	.35	.11	.35	1.00								
IV. Paragraph	.55	.49	.29	.43	.37	1.00							
V. Organization	.45	.42	.26	.38	.32	.39	1.00						
VI. Para. Analysis	.53	.53	.36	.47	.35	.43	.43	1.00					
A. Main Idea[1]	.30	.35	.24	.36	.12	.23	.12	.72	1.00				
B. Develop-ment[1]	.40	.40	.26	.33	.30	.35	.37	.74	.40	1.00			
C. Organization[1]	.39	.33	.32	.35	.35	.32	.33	.75	.44	.34	1.00		
D. Supt. Detail[1]	.47	.51	.21	.33	.24	.33	.42	.71	.30	.38	.44	1.00	
Objective Total Test	.65	.61	.53	.66	.55	.68	.68	.85	.54	.66	.67	.63	1.00

1 = Items within Paragraph Analysis

ATTACHMENT 9

Regression of Rating Scales on Objective Measure

REGRESSION OF ANALYTIC SCALE TOTAL ON SIX OBJECTIVE MEASURE SUBTESTS

Objective Measure Subtests	Unstandardized Coefficient	Standard Error of β	Standardized Coefficient	F
Main Idea	.141	1.300	.009	.012
Development	.428	1.595	.025	.072
Supporting Detail	1.502	1.669	.082	.811
Paragraphing	4.037	1.514	.260	7.105*
Organization	1.771	1.173	.142	2.277
Paragraph Analysis	2.184	.674	.330	10.488*

df = 1, 92

* = $p < .01$

R = .66

r^2 = .40

REGRESSION OF IMPRESSIONISTIC SCALE TOTAL ON SIX OBJECTIVE MEASURE SUBTESTS

Objective Measure Subtests	Unstandardized Coefficient	Standard Error of β	Standardized Coefficient	F
Main Idea	.025	.274	.007	.008
Development	.497	.336	.137	2.182
Supporting Detail	.209	.352	.051	.354
Paragraphing	1.06	.319	.310	11.088*
Organization	.421	.248	.153	2.894
Paragraph Analysis	.371	.142	.254	6.08*

df = 1, 92

* = $p < .01$

R = .68

r^2 = .46

ATTACHMENT 10

Regression of Analytic Rating Subscales on Groups of Objective Items

REGRESSION OF ANALYTIC SUBSCALES ON ANALOGOUS OBJECTIVE MEASURE SUBTESTS

Analytic Subscale	Objective Subtest	Unstandardized Coefficients	Standard Error of β	Standardized Coefficients	F	df	R	r^2
I. Focus	Main Idea	.518	.269	.188	3.709*	1,96	.37	.14
	Main Idea[1]	1.051	.375	.273	7.826*	1,96		
II. Development	Development	.660	.307	.213	4.619*	1,96	.40	.17
	Development[1]	.953	.333	.283	8.171*	1,96		
III. Support	Supporting Detail	.545	.313	.161	3.026*	1,96	.48	.23
	Supporting Detail[1]	1.548	.347	.412	19.896*			
IV. Paragraphing	Paragraphing	1.303	.269	.442	23.511*	1,97	.44	.20
V. Organization	Organization	.905	.360	.360	13.516*	1,96	.42	.18
	Organization[1]	.514	.131	.131	1.797	1,96		

1 = items within Subtest VI, Paragraph Analysis

* = $p < .01$

ATTACHMENT 11

Means and Standard Deviations on the Analytic Rating Scale, Impressionistic Rating Scale and Objective Measure

Measure	Mean	SD	Average Percent Correct	N
I. ANALYTIC SCALE				
Subscales				
1. Focus (Main Idea)	10.58	3.13	66%	112
2. Development	10.71	3.19	67%	112
3. Organization	9.91	3.37	62%	112
4. Support	10.33	3.07	65%	112
5. Paragraphing	10.64	3.21	67%	112
6. Mechanics	10.60	2.63	66%	112
TOTAL SCALE	62.77	16.78	65%	112
II. IMPRESSIONISTIC SCALE				
Total Scale	13.33	3.72	56%	112
III. OBJECTIVE MEASURE				
Subtests				
1. Main Idea	3.56	1.13	71%	113
2. Development	3.54	1.05	71%	114
3. Supporting Detail	4.63	.88	93%	114
4. Paragraphing	4.23	1.09	83%	115
5. Organization	3.93	1.38	79%	116
6. Paragraph Analysis	8.21	2.56	69%	114
TOTAL TEST	28.39	5.45	76%	109